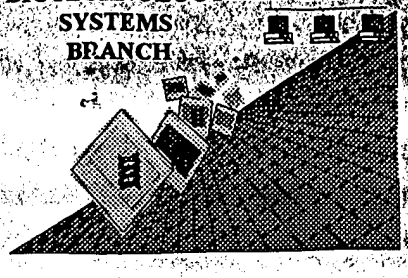
BIOTECHNOLOGY SYSTEMS BRANCH

RAW SEQUENCE LISTING ERROR REPORT



The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/847,513

Source: OIPE

Date Processed by STIC: 5/16/2001

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY
FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)
PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO **REDUCE** ERRORED SEQUENCE LISTINGS, **PLEASE** USE THE **CHECKER VERSION 3.0 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

Checker Version 3.0

The Checker Version 3.0 application is a state-of the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.
Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO). Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

Checker Version 3.0 can be down loaded from the USPTO website at the following address:
http://www.uspto.gov/web/offices/pac/checker

Raw Sequence Listing Error Summary

ERROR DETECTED **SUGGESTED CORRECTION** **SERIAL NUMBER:** 09/847,513

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

1 ____	Wrapped Nucleics	The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3, as this will prevent "wrapping".
2 ____	Wrapped Aminos	The amino acid number/text at the end of each line "wrapped " down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3, as this will prevent "wrapping".
3 ____	Incorrect Line Length	The rules require that a line not exceed 72 characters in length. This includes spaces.
4 ____	Misaligned Amino Acid Numbering	The numbering under each 5th amino acid is misaligned. This may be caused by the use of tabs between the numbering. It is recommended to delete any tabs and use spacing between the numbers.
5 ____	Non-ASCII	This file was not saved in ASCII (DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text so that it can be processed.
6 ____	Variable Length	Sequence(s) ____ contain n's or Xaa's which represented more than one residue. As per the rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the (ix) feature section that some may be missing.
7 ____	PatentIn ver. 2.0 "bug"	A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequence(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. **This applies primarily to the mandatory <220>-<223> sections for Artificial or Unknown sequences.**
8 ____	Skipped Sequences (OLD RULES)	Sequence(s) ____ missing. If intentional, please use the following format for each skipped sequence: **(2) INFORMATION FOR SEQ ID NO:X:** **(i) SEQUENCE CHARACTERISTICS:**(Do not insert any headings under "SEQUENCE CHARACTERISTICS") **(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X:** **This sequence is intentionally skipped** Please also adjust the "(iii) NUMBER OF SEQUENCES:" response to include the skipped sequence(s).
9 ____	Skipped Sequences (NEW RULES)	Sequence(s) ____ missing. If intentional, please use the following format for each skipped sequence. **<210> sequence id number** **<400> sequence id number** **000**
10 ____	Use of **n's** or **Xaa's** (NEW RULES)	Use of n's and/or Xaa's have been detected in the Sequence Listing. Use of <220> to <223> is MANDATORY if n's or Xaa's are present. In <220> to <223> section, please explain location of **n** or **Xaa**, and which residue **n** or **Xaa** represents.
11 ____	Use of "Artificial" (NEW RULES)	Use of "Artificial" only as "<213> Organism" response is incomplete, per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.
12 ____	Use of <220>Feature (NEW RULES)	Sequence(s) ____ are missing the <220>Feature and associated headings. Use of <220> to <223> is MANDATORY if <213>ORGANISM is "Artificial Sequence" or "Unknown" **Please explain source of genetic material in <220> to <223> section.** (See "Federal Register," 6/01/98, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of new Rules)
13 ____	PatentIn ver. 2.0 "bug"	**Please do not use "Copy to Disk" function of PatentIn version 2.0.** This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other means to copy file to floppy disk.

AMC - Biotechnology Systems Branch - 4/06/2001

OIPE

RAW SEQUENCE LISTING DATE: 05/16/2001
PATENT APPLICATION: US/09/847,513 TIME: 13:37:02

Input Set : A:\MBA10101.txt
Output Set: N:\CRF3\05162001\I847513.raw

Does Not Comply
Corrected Diskette Needed

pp. 2-16

3 <110> APPLICANT: MBARI
4 DeLong, Edward
5 Beja, Oded
7 <120> TITLE OF INVENTION: Light-driven energy generation using proteorhodopsin
9 <130> FILE REFERENCE: MBA-101
C--> 11 **<140> CURRENT APPLICATION NUMBER: US/09/847,513**
C--> 11 **<141> CURRENT FILING DATE: 2001-05-01**
11 <150> PRIOR APPLICATION NUMBER: 60/201,602
12 <151> PRIOR FILING DATE: 2000-05-03
14 <160> NUMBER OF SEQ ID NOS: 65
16 <170> SOFTWARE: PatentIn version 3.0
18 <210> SEQ ID NO: 1
19 <211> LENGTH: 105184
20 <212> TYPE: DNA
21 <213> ORGANISM: Naturally occurring gamma proteobacterium
23 <220> FEATURE:
W--> 24 **<221> NAME/KEY: CDS(complement)**
25 <222> LOCATION: (50866)..(51615)
26 <223> OTHER INFORMATION: light-driven proton pump; has the properties of a light-driven pr
27 oton pump when expressed with retinal in Escherichia col
30 <300> PUBLICATION INFORMATION:
31 <301> AUTHORS: Beja,O., Aravind,L., Koonin,E.V., Suzuki,M.T., Hadd,A.,Nguyen,L.P.,
32 Jovanovich,S.B., Gates,C.M., Feldman,R.A., DeLong,E.F
33 <302> TITLE: Bacterial rhodopsin: evidence for a new type of phototrophy in the sea
34 <303> JOURNAL: Science
35 <304> VOLUME: 289
36 <305> ISSUE: 5486
37 <306> PAGES: 1902-1906
38 <307> DATE: 2000-09-15
39 <308> DATABASE ACCESSION NO: AF279106
40 <309> DATABASE ENTRY DATE: 2000-06-15
41 <313> RELEVANT RESIDUES: (50866)..(51615)
43 <400> SEQUENCE: 1
44 ttgttatatc agtaatggct attgctccaa taacttaata ctaatatata attagtttat 60
46 gaataaattt tatatatttg ggttattgtt ttttacacta aatgcatttt cttgctcaga 120
48 tcttctagat acagacatga gagttcttga ttccgctgag tcaagaaacc tttgcgagtt 180
50 tgaaggaaaa gctttactag ttgtgaatgt tgcaagtaga tgtggttaca cttatcaata 240
52 tgctggcctt caaaagttat atgaaagtta taaagatgaa gattttctag taattgggat 300
54 cccatctaga gattttcttc aagaatactc tgatgaaagc gatgttgcag aattttgttc 360
56 tacagaatac ggtgttgaat ttcctatgtt ctcaactgct aaagtcaaag gaaaaaaagc 420
58 acacccattt tataaaaaac ttattgcaga atcaggtttt actccctcat ggaactttaa 480
60 taaatactta atctcaaaag agggcaaggt tgtatccaca tatggatcaa aggtaaagcc 540
62 tgattcaaaa gagcttatat cagctataga aggcttgctg taaaattatt acttagaaac 600
64 taatacagtt ttaggcttgt ttgctgcaaa tattccatta tctacaactc caggaatatt 660
66 attaatcaaa gcttccattt cagtggggtt tgaaatatcc atattagaga tatctaaaat 720
68 gtgattacct tggtctgtta taaatccagt tctatatgtg ggtattccac cgatcgagat 780
70 tatttttctt gcaacaaggc tcctactttc aggtatcacc tctataggca gtggaaaagc 840

RAW SEQUENCE LISTING DATE: 05/16/2001
PATENT APPLICATION: US/09/847,513 TIME: 13:37:02

Input Set : A:\MBA10101.txt
Output Set: N:\CRF3\05162001\I847513.raw

```
  72 tcccaaaaga ttaaccatct ttgactgatc aactatacat ataaactcgt tagaggcaga    900
  74 agcaactatc ttttctctag tatgtgcgcc accaccacct ttaataagac aattttcagg    960
  76 agacacctca tctgcaccat ctatgtaata agctatatca actacatcat taaggctaaa   1020
  78 gacctctatc ccattttcat ttaataattt tgatgaagca tctgaactag aaacagctcc   1080
  80 agcaaatttg tgcctatgct cctttagttc ttctataaaa aaattaactg ttgagccggt   1140
  82 tccaatacct aaaatcatct caggatgaag attatttttg atatattcta tagcttgttt   1200
  84 agcaacattt atctttgagc cactcataga gttataatac aagaaaatat aggtagttaa   1260
  86 ttattttgag actaaaaatt aaaaaaacag gttcttttaa gaattcccag aagtacctaa   1320
  88 agcttattcc taatgcatct gtttatgacg tagcattaaa atcacctata acatttgctc   1380
  90 taaatatttc ttcaaagctg gggaataaag ttttcctaaa aagagaggat ctgcaaccta   1440
  92 tattttcttt taaaaacaga ggagcgtata acaagattgt aaatttatcc gatgccgaaa   1500
  94 agaagagggg ggttattgct gcatcagcag gaaatcatgc tcaaggggta gccagtgcat   1560
  96 gtaagaaatt aaaaattaat tgcttgatag ttatgccaat aacaactcca gaaataaaaa   1620
  98 taaaagatgt aaaaagattt ggagccaaaa tactccaaca tggggacaac gtagatgcag   1680
 100 cattaaaaga ggcactgttt attgcaaaga aaaaaaaatt gtcttttgtt catccttttg   1740
 102 acgaccctct aacaattgct ggccaaggga ctataggaca agaaattctt gaagataaaa   1800
 104 ataattttga tgttgtcttt gttccggtgg gaggaggagg tattctagct ggtgtatctg   1860
 106 cctggatagc acagaataat aagaaaataa aaattgttgg tgttgaggtt gaggattccg   1920
 108 cttgtcttgc tgaggccgta aaagctaata aaagagttat tttaaaagaa gtgggcctct   1980
 110 ttgctgatgg ggtggcagta tcaagggttg gaaaaaataa ttttgatgtt attaaagagt   2040
 112 gcgtagatga agtcattaca gttagcgttg atgaggtctg caccgctgta aaagatatct   2100
 114 ttgaagatac aagggttcta tcagaacctg ctggggcatt agcacttgca gggttaaaag   2160
 116 cctacgcaag gaaagttaaa aataaaaaac ttattgctat aagttctggc gctaatgtaa   2220
 118 atttccaaag acttaatttt attgttgagc gatcagagat tggtgaaaat agagaaaaaa   2280
 120 tattaagtat caaaatccca gagatacctg gaagttttct taagctttca aggatgtttg   2340
 122 gcagctctca agttacagag tttaactaca ggaaatctag cttaagcgat gcatatgttt   2400
 124 tagttggtgt tagaactaaa actgaaaaat catttgaaat cttaaagtcc aaattaaaaa   2460
 126 aagcaggctt caccttagc gactttactc gaaatgaaat atccaatgat catctgaggc   2520
 128 atatggttgg tggcagaaat agtgactcag gctctcataa caatgaaaga atatttaggg   2580
 130 gagagtttcc tgagaagccg ggcgcgctgt taaattttct agagaaattt ggaaataaat   2640
 132 ggmatatttc cttatttcat tacaggaacc taggttcagc ttttggaaag atattaattg   2700
W--> 134 gcatcgagag taaggataaa gacaagctaa taaatcattt aaataagtca ggnactattt   2760
 136 ttacagaaga aacctctaac aaggcataca aagatttttt aaaatgaaag gttaatactt   2820
 138 taatctaaat ttaattgaaa aaagctcatc gctagggttt tcccacggct ctttgaacaa   2880
 140 ctcggattga gatctatcat cctcctcgtc gtaaattctc ccacctttag aatagaccaa   2940
 142 aaatagatat gacaaaggag cgagctcata tttatatcta atttggaacg aagctacgcc   3000
 144 agtattaaat tcattaacta tattatttcc tttataaagg tatccattag catctgaaaa   3060
 146 aatactaata ggatttttg ctttcaaagc aacaaattga ctcttaagtc tgatctcatg   3120
 148 tttattattt ttaaaccaat ttagatcaaa agataaggta tcttgtcttg aatcatatga   3180
 150 ggcaagatta ttattatcct gccatatcag ccattcattt tcctttctta ttctatattg   3240
 152 cgcattgatt cttaagttat catttggaaa tattgaacct gctatcttgt aaaatttttct   3300
 154 accataccca ttcgaatccc accgattatc tttctctccc ttaaagaagc taactctcca   3360
 156 gtcatatgtc cagaatgaat agttctttgc ctcaaagtct gctgtaatac ctattcttct   3420
 158 ttttgatttg ataaaggggt aggcttcatt ttttcttgtg atagttgtat ttttcccaga   3480
 160 agatctaaag ttaaaatcta attgaaattt agagttgtcc ttaaaactaa aagaattttt   3540
 162 ttgatcgatg cctattggat ttgaattacc agctgtgtca gcatcataat ttagatcaat   3600
 164 tccataatct atttgtttta atatcgagct attatcaaat tcatttattt ttcgatttcc   3660
 166 tccaatacca gcatgaatcc agtctcttct ttgcagataa ccaaagtcat ttaactcgaa   3720
 168 gtcgtcttca aaataaagaa ggcttccact tatgttagat agtttatttg gaagatatgt   3780
```

see item 10 on Error Summary Sheet

RAW SEQUENCE LISTING DATE: 05/16/2001
PATENT APPLICATION: US/09/847,513 TIME: 13:37:02

Input Set : A:\MBA10101.txt
Output Set: N:\CRF3\05162001\I847513.raw

170 aaactgagtc ctatacccaa gcccattttt accatctttt tctgaggcta acagatctga 3840
172 atatgtaatt aatttttttg aacgaatatt gatgtaatca ataacattga ctgttgatga 3900
174 ttcgcctgtc atttcattct caacattcgt taccatgaag ccaagcgttt tatttccaag 3960
176 ctttgttcga gatcgaagag cataatagtc tcttccaact gaaaaggctt catcagcctc 4020
178 acttgctaca aatactccaa attcattatt attacttttt tgagtaagtc ttaatgcaaa 4080
180 atcaatatca gaatagtttt ttttagctgc ctcgcaacct tcttcattac tttcttctga 4140
182 gcaattatag ctgggggcag ctccaatcct ccgcgtattt ataaccgagt atctatcata 4200
184 attactaata tcaaatagtg attggttttc attgaaaaat gctctttttt ctgagtaaaa 4260
186 agtttcttga gcagaaaagt taataaccac atcatcactc tcagcttgtc cgaaatctgg 4320
188 attaatagct aaatttattt gacgaccttt tccagtgcta taaaagattt cagccccaat 4380
190 atctgaccct tcttggtttg taactgaatt tttatttgaa gatatatatg gaaaaaaagt 4440
192 aagctttgat tttgtatagt tttgtatttc taagctatct aactcttgaa agtagtcatt 4500
194 tctactagcc attgttccgg cactgctaac ccatgactca ttttcggcac tataacgtaa 4560
196 tgcggtgtaa ttaatttttc ttatatcacc atcaggctgt ttcattaacg ttacatccca 4620
198 aggaataaaa aactcagaga cccaataccc atcaaatttt tgtgtttttg caatccaatc 4680
200 tccatcccag tctgtcttaa agtctcctgc ttgcgttttt atggcatcga aaagcgagtt 4740
202 cccaagattt atagcaagaa tgaaagcttt gttaccatca ccatcaaagt ctatatttat 4800
204 agagttttta tcgctaagtg aatttatttg atctctaagc gtcttcctgg agaacataga 4860
206 atcattactt tgaaaattct taaatccaac atatatacca tccttatttg agaaaattaa 4920
208 agccgttgta agaagttcat ttttcttaag agtaaaagga gatgtttcat aaaaatctgt 4980
210 aatttcaaat gcattattcc actcaggctc atcaagagag ccatcaataa caattgagtt 5040
212 tgaccaaatc agcacagagg taagtaaaag tgataatgag gctaaagtta ttttcataga 5100
214 tagatttaat tcagagtatt ttaatctatg aagagtagga aatcatccga tcattctcag 5160
216 aaaaaacata atcatattta aggtcatgtt tttctccaaa atgcatatta caaatctggt 5220
218 attgataaca cagtccaacc aataaaggtc ttgatacctc ctcgtttata gagccaataa 5280
220 ttctatcgaa atatccagag ccatagccaa gcctgtatcc atttaaatca actcctgtca 5340
222 taggaataaa cattaaatca atttcattta tgttgacata atcttcactt ttaacctctt 5400
224 tgatcccaaa ttgatttata aagaagttag gctgctcatc cagcaaatta aaagccatca 5460
226 tttcgtcatc aattactttt ggaatataaa tattcttttt aagtttagta aaagcttgaa 5520
228 ttaacaaatt tgtattgact tcatttcgaa aaggaaaata taaggcaatg ttttgcattt 5580
230 catgagtatt gatttttttct aatacatttt cttggattaa aaaacttata ttgctcttag 5640
232 ataaatcaga aatagactgc ccctgctcaa agagtgattt tcgtatttta tttttcacca 5700
234 taaattgtgc cgaggctaac aaaccaaatc accgccatgt gctatacctg aaccgatggg 5760
236 tcaggtggtg aatttcataa caaatcaggc ttccctgtaa aggtactgcg caacaatgct 5820
238 agaggaatat cacctattta attgtatcgg ttcaaattta ttaacacatt agcgtatgaa 5880
240 ccagaatgca gttaattata aaatatatag gtattaagta aaagttaatt tttagagagc 5940
242 agactctatt ttttgtatta gcttttcgat atctttatta tttaaagcag agtcagtatt 6000
244 ctcggttggc ttggataaaa gcttactagc catagttaga ccagctatca ccaaggcatt 6060
246 attcttatca ctgatgccat caagctcatc atttaataat tttgcagctc taatgagttg 6120
248 gtctttctct tctggcggac aggctaacgt cagatctcgg ccaaaaattc ttaaagatag 6180
250 cgtttccatt tctgacatta ttttttgagt ttagttattt ctttatttag tatgagctta 6240
252 tctttttttcc agtctctttc cctttgcaca taagaatcta ttattccttt ttgctcatca 6300
254 aatcttttta tcaaaagatc taccttatcc tctagttcaa ataatgatga tttttttttct 6360
256 tctgacattg tataagttta attgtactag aatgaattgg aaagtttctt taataaaaga 6420
258 gtaaaataag gcatggaaaa aataatttttt aaaaacagaa gagattcact tataaaacac 6480
260 ctacctaaga attcagcctt aattgtacct ggtgcagatt tgcaatatag aaatgctgac 6540
262 tcatcttata atttaagaca ggaaagtagc ttctactatc tgtctggctt ttgtgagccc 6600
264 tcttctctaa tggttttagt taataatgga aaaagcattg attcaataat ttttgttcct 6660
266 gaaaaagata aacttaaaga aatctgggat ggttatcgag ctggccctga gggcgcaata 6720

RAW SEQUENCE LISTING DATE: 05/16/2001
PATENT APPLICATION: US/09/847,513 TIME: 13:37:02

Input Set : A:\MBA10101.txt
Output Set: N:\CRF3\05162001\I847513.raw

268 aatgattttc tttttgatca agcttttgaa aataataaat cagatgcttt aatgcctgaa 6780
270 atccttcaag ggctagaaaa agtttttttat tcaataggga agaaaaatgg ctttgatcag 6840
272 aaagtaattg actggacatg cgcagcaaat tctaaagata ggcacagcaa atcaattgat 6900
274 attattgatg gctcttcgat ggtaggaaat ttaaggctta tcaaagataa gcatgaaatt 6960
276 gatattatga agagagcttg tgaaatttca gctgaatcat atattgaggt catgaaatct 7020
278 ataaagcctg gggacaatga gcaggaaata gaggcgctat ttttatatga attcgccaaa 7080
280 aggggaggaa ggtttccagc ttatacacct atagttgctg gaggtgaggg tgcttgtgta 7140
282 ttgcattata ttgaaaatga taaagagtta gcttcatcag atttaatttt ggtagacgca 7200
284 ggatgtgaat acaaaatgta tgcatctgat atcacaagaa ccttcccagt aagtggaaaa 7260
286 ttttcagatg aacagctaca aatttataat attgtccaca aagccaatct tgctgcaatc 7320
288 gatgctgtaa aaactggtaa tagcataatg gagccccaaa tggtttcaga aaaagtaatt 7380
290 actgaaggtc ttgtagagtt gggtattcta tctggcgatg ttaatcagct tcataaaaat 7440
292 ggtgcattca aggactttta tatgcataag gtgggacatt ggcttggact tgatgttcat 7500
294 gatgttggtg actacatgga gggagatgag tttatgaagt ttaagccagg gatgataacc 7560
296 acaatcgagc caggcatcta tatcagtagc gcaatggatg tagatgacaa atggaaaggc 7620
298 atcggcataa gaatagagga cgacatcctt gtaacagatt caggcaatat taatctaaca 7680
300 gagaaggtgc catctaatcc tcaagaaata gaatcattga tggcttagac tatggaggtt 7740
302 ccaattgtta tttctggcgg agggataata ggtaattaca tttctcttag gcttgaaaaa 7800
304 aataatatca aaaccgttat tgtcgaaaaa gctagtagtt tcaaagccct agataagggt 7860
306 ataagaacag tcactctcaa tgagcattct atgcaaatgc taaaaaatat tggtatttgc 7920
308 ccatcaattg ctcaaatcaa cagcatcgac gtattagatg gtgagggtac aggcaaaatt 7980
310 caatttctag caaaggacgt aggcagcgaa aacctttcat atgtaaccta tttcaatgaa 8040
312 ttacaaaaac taatttctga tccatgtaaa gaaagaacct tatttaataa tgagattgat 8100
314 tcagttcaga atcttaatac agaatctgat ccagagatca tgcttaaaga tggcatgacc 8160
316 ataaaaacga atctaattgc tggatgtgat ggaagaaatt caaatattgc aaaaattgct 8220
318 tcacttacaa gcagcttcga tgactactta caaacagctt taacttttgt cgttgatatt 8280
320 gataatgatt cacatggcaa agctcaccaa gtttttttctg aaaaaggaat atttgcactt 8340
322 atgccactcc cagaaggcaa gggtgagatg aataaatgca cagtggtttg gtcaataaaa 8400
324 aatcaagttt tgggagatga gcctgtatct gagtttgtaa aaaataacat ttctttttt 8460
326 gaatcaaagc ttaatgttag tctcagggtt aagtcagaaa ttttaagttt taaattatcg 8520
328 aaccatcatt ttgaaaacta tattagcgga cctattgttc ttcttggtga tgctgctcac 8580
330 tcaattcacc ccttagcagg tcaaggtatt aatctaggat ttgcagatgc agatactttt 8640
332 tgtgaagagg taattagttc ttataaaaaa gggattgcct ttaatgagaa atcagtttta 8700
334 aaaagatatg agattagaag aaaaagtatg aacttttttaa tgttgaagtc tatggacttt 8760
336 tttgtggatt tatttggttc agaaaattta tatctaaggc tgatcagaaa tttaggtatt 8820
338 tcttcgttaa ataaatcaaa gtttgttaaa gcattcttta taagacatgc ctctggaatg 8880
340 aataagtttt aaatttgtat taaacttttt gacctttagc tctaagttct ttaagaactt 8940
342 cactaatgcc ttttttatca atgattctca tacctttttgc agatacttta agatttacga 9000
344 acctgttctc agattcaacc caaaatttgt gtgtgtgaag attaggaaaa aactttcttt 9060
346 tagtcctatt tttagcgtga gaaacattgt ttcctgactg tggtatctta cctgttactt 9120
348 gacatatttt actcattgaa acgcgatttt atagaacact gaggaactta gcaatactat 9180
350 tgtgaaacaa atttatttat tacggcatgc acaatctgac tgggagagct ctaatcagaa 9240
352 agattttgat agaccattag caagaaaagg cattgaagaa gcaaataaaa tatcatgtta 9300
354 ctgcaaatct cattcaattt tagtagataa aatattctgt agcactgcag aaagaactaa 9360
356 gcagactttt gatatatgca gtgatgggct taattatcca atagctgaag cagtctatac 9420
358 tgatgagctt tacttttctg gccctggtga gatcgttaag cttatccaaa gtttaagtga 9480
360 attcatttcc tctgttttaa taataggcca caatccatca atgcaaatgt acatagatgc 9540
362 tatttcagaa aatcctcata ttacgtattc aacatgcggg ctggcagaaa ttctcgttga 9600
364 aagttcatgg aaagacttat ctttaaaaaa atgtaagtta aaatctttta ttcaaccagg 9660

RAW SEQUENCE LISTING DATE: 05/16/2001
PATENT APPLICATION: US/09/847,513 TIME: 13:37:02

Input Set : A:\MBA10101.txt
Output Set: N:\CRF3\05162001\I847513.raw

366 agagctttaa aaattgaaaa acataaaaat taagatcatt aacccactaa tgggatccaa 9720
368 gataccctta cctcaatatg aaacaaaggg ctcggcagga ttggatttaa gggcatgcct 9780
370 agatagtaat ctcagccttc aagcaggaac atctcagttg atacctattg gttttgcaat 9840
372 gtacttagaa gatcccggtc ttgcagcaat ggttatacct agatcaggtt taggttctaa 9900
374 gcatggaatc gttcttggta atctggttgg gttgattgat tcagactatc aaggagagct 9960
376 aatggttcct gcctggaata gatcagatac agattttgag attaatcctg gagacaggat 10020
378 tgcacaaatg attatagttc cagtgattca agcagatttt gaaattgtag acgagttcaa 10080
380 tgagactcag aggggagaaa agggttttgg aagttcaggt ataaattgat aaatttactt 10140
382 tttcttgcca aatctttctt caaatttctg aactcttccg ccagtatcaa taatttttg 10200
384 cttaccggta taaaaaggat gagaagcaga ggatatatca agagggtagt atgggtatgt 10260
386 ttttccatct tcccattctt tcgtttgagt cgtatctaat gttgaacgaa tgagaaagaa 10320
388 cttatcagca ctagcgtcat ggaataaaac ttcacggtat tcaggatgta tatcttttt 10380
390 cataataaaa tcaaaatttg gatgagaact ataacaaaaa acaacttctt ttcaatcaaa 10440
392 acatgaaaat tttttactat gacatagctg tttcattgcc tctaaggcaa tgctttactt 10500
394 atagttctga gcttaaaatt acaaaaggaa cccgcgtgtc agttcctttt ggtaaaagga 10560
396 agattgtagg agtggttata aaaaatattc aaaagccaga tttcctaaaa aaagcgggag 10620
398 ctataaaaaa aattatcgct gtacttgatg aatatccttt gttcgacaag ccaatttttg 10680
400 attctatatt gtggtcttct gattactatc atcatcctat tggtgaggtt tttaatacct 10740
402 ttatacccac cgaattaaga aaaattaata ataaaaaaat tgaagcttta agagaatttt 10800
404 ctgaatattc agtaaatgag gatgataaga aattcgattt aaccaaggat caagaaaaag 10860
406 cagtcaaggc gctttctaaa tctaaaggat tttcacccac tttattatat ggagttacag 10920
408 ggtctggaaa aacagaagtt tacttaagag ttgccgaaac ttttattaaa aataataagt 10980
410 cagtattagt tttagttcca gaaataaatt taacccccca attgctctcg cgatttgaga 11040
412 atagatttaa tggtgaaatt ggcatatatc attctaagca aacagcagct aagagattaa 11100
414 agacttggct aaaagctaaa tttggttcta taaaaataat agtaggaact cgatcttcgg 11160
416 ctttagtgcc tttagataac attggtttaa taattatcga tgaagagcat gaccaatcat 11220
418 ttaggcagtc agaagggttt aaattctctg ctagagactt aagtataaaa agggcacagc 11280
420 ttgcagatat tccaattatt ttgggatcag caacccccttc gctgcaaact ttaaaacttg 11340
422 taaaagaaaa taaatttata agagttgata ttcctaatcg agttgatgga aacaagcctc 11400
424 ctaaattaat agccttagat atcaataaca gcccctttaat aggcggagtt gctaaagaga 11460
426 caattgaagc aatgcaatca accatagaca gaggagaaca ggttctagtt tttattaata 11520
428 gacgaggatt cgctccactc tatcaatgta gtagttgtgg gtgggtagca gattgtaaat 11580
430 cttgtgatac aaatttagtc ttccaccagg caagaaatag attaatttgt cataggtgtg 11640
432 aatctgccta ctctgttaat ttgtcttgtc cggcatgcaa gtctaatgac tttaatatgt 11700
434 atggagctgg aacagagaga gttgaagaag ttcttaaaag cagctttgta aagactccaa 11760
436 taattagagt tgatcatgac tcaacaaaaa aagtgggagc tatggaggct atagttaaaa 11820
438 aaattcattc ctcagacgca gcaattttag ttggaactca aatgcttgca aaaggacatg 11880
440 attttcctaa agtcacctta agcgttattt taaatgctga taatggcctt ataagcccag 11940
442 aaattaatgc attagagaaa atatctcaat tgcttattca ggtctctgga agagcaggaa 12000
444 gaaataataa tcttgcaaaa gttattattc aaacaagata tcctgatgat ataaatctta 12060
446 ataaaattaa gacaggagat tatatgaaat ttgcttctca atgcctaagt accaatgagc 12120
448 aaatgaactt acctccattt actactttat gtctgcttag gtgctcatca ccaactcaaa 12180
450 agagtaatgt agatttccta gagaaagctg ttttaatttt atccaatagg actgatataa 12240
452 atgttattgg tcccttgcct tcattagttt cgaagtcgaa aggaaattat aggcaccaag 12300
454 tctatatcca tgcaccaaag aagacttttt taaataaggt attaaagttt ttgacaacag 12360
456 agtttgaaaa atggccggaa tctaataagg ttaagtggtc tttcgacatt gatccaatag 12420
458 acttaagcta aatattaatc ttaattaatt gtcctgggta tattggttta ttgtttagtt 12480
460 tattctctgt attaatttct tctacagtca ccccaaatct tatcgctatt tctgataaga 12540
462 catccccttt ttgtattttg taagtcacaa agcctggatc aatactcata aaggtatttg 12600

<210> 4
<211> 750
<212> DNA
<213> Naturally occurring gamma proteobacterium

<220>
<221> CDS
<222> (1)..(750)
<223> light-driven proton pump; has the properties of a light-driven pr
oton pump when expressed with retinal in Escherichia coli. Note
that additional three nucleotide residues incorporated by pcr pri
ming with reference to the original 31A08 DNA sequence (DNA resid
ues 4-6, ggt), adding a new restriction site for cloning

FYI: Per 1.823 of Sequence Rules, the <2237 response
has a MAXIMUM of 4 lines.

VERIFICATION SUMMARY DATE: 05/16/2001
PATENT APPLICATION: US/09/847,513 TIME: 13:37:03

Input Set : A:\MBA10101.txt
Output Set: N:\CRF3\05162001\I847513.raw

L:11 M:270 C: Current Application Number differs, Replaced Current Application No
L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:24 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:1
L:134 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:3559 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:2
L:3572 M:286 W: Invalid Database Entry: Valid Formats YYYY-MM-DD or Mon-YYYY, SEQ ID:2
L:3584 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:3
L:3597 M:286 W: Invalid Database Entry: Valid Formats YYYY-MM-DD or Mon-YYYY, SEQ ID:3
L:3615 M:259 W: Allowed number of lines exceeded, <223> Other Information:
L:3790 M:286 W: Invalid Database Entry: Valid Formats YYYY-MM-DD or Mon-YYYY, SEQ ID:6